The quasi-Gaussian entropy theory

Amadei, A.; Apol, M. E. F.; Di Nola, A.; Berendsen, H. J. C.

Published in:
Journal of Chemical Physics

DOI:
10.1063/1.470744

The quasi-Gaussian entropy theory: Free energy calculations based on the potential energy distribution function

A. Amadei, M. E. F. Apol, A. Di Nola, and H. J. C. Berendsen

Citation: The Journal of Chemical Physics **104**, 1560 (1996); doi: 10.1063/1.470744
View online: https://doi.org/10.1063/1.470744
View Table of Contents: http://aip.scitation.org/toc/jcp/104/4
Published by the American Institute of Physics

The quasi-Gaussian entropy theory: Free energy calculations based on the potential energy distribution function

A. Amadei, M. E. F. Apol, A. Di Nola,[a] and H. J. C. Berendsen
Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

(Received 7 August 1995; accepted 18 October 1995)

A new theory is presented for calculating the Helmholtz free energy based on the potential energy distribution function. The usual expressions of free energy, internal energy and entropy involving the partition function are rephrased in terms of the potential energy distribution function, which must be a near Gaussian one, according to the central limit theorem. We obtained expressions for the free energy and entropy with respect to the ideal gas, in terms of the potential energy moments. These can be linked to the average potential energy and its derivatives in temperature. Using thermodynamical relationships we also produce a general differential equation for the free energy as a function of temperature at fixed volume. In this paper we investigate possible exact and approximated solutions. The method was tested on a theoretical model for a solid (classical harmonic solid) and some experimental liquids. The harmonic solid has an energy distribution, which can be derived exactly from the theory. Experimental free energies of water and methanol could be reproduced very well over a temperature range of more than 300 K. For water, where the appropriate experimental data were available, also the energy and heat capacity could be reproduced very well. © *1996 American Institute of Physics.* [S0021-9606(96)51004-8]

I. INTRODUCTION

Free energy is the quantity of central interest in physics and chemistry, since the free energy determines the equilibrium physics, and free energy differences determine chemical equilibrium and dynamics.

Earlier attempts to calculate the Helmholtz free energy and to relate it to well measurable basic properties were restricted to either diluted gases[1] or gases and monoatomic van der Waals liquids at relatively high temperature, using the so-called High Temperature Expansion of Zwanzig.[2] This last expression, applied to the free energy with respect to the hard sphere liquid, is given by[3]

$$A^*=\langle \nu\rangle_0-\frac{\langle(\nu-\langle \nu\rangle_0)^2\rangle_0}{2\,kT}+\frac{\langle(\nu-\langle \nu\rangle_0)^3\rangle_0}{6\,(kT)^2}-\cdots \qquad (1)$$

with ν the perturbation potential energy, where the ensemble averages $\langle\cdots\rangle_0$ are taken over the reference state. Often Eq. 1 is truncated after the second term and in this case the right hand side can be expressed in terms of $g(r)$ or related quantities, using either the Barker–Henderson theory[4,5] or the Chandler–Weeks–Andersen theory.[5,6] Later extensions to small molecular fluids are reviewed by Gray and Gubbins.[7]

In this paper we will rephrase the basic equations of statistical mechanics in terms of the one-dimensional distribution function of the total potential energy, instead of the multi-dimensional partition function. Since we can define the internal energy and the Helmholtz free energy in terms of this energy distribution function, all other thermodynamic properties can be derived too.

From the fact that at not too low temperature a macroscopic system can be regarded as an infinite collection of identical independent subsystems, it follows that the distribution must be close to a Gaussian (we will refer to these macroscopic systems as quasi-Gaussian systems). Hence for every quasi-Gaussian system it is possible to classify different ''statistical'' states, according to the corresponding uninormal distribution, and to derive a general expression of the reduced entropy (the quasi-Gaussian entropy) in terms of the heat capacity and a dimensionless and intensive quantity, the intrinsic entropy function.

The article is organized as follows. In Sec. II A we define the reduced and ideal reduced free energy, and in Sec. II B we give the general shape of the potential energy distribution function and the physical and mathematical restrictions, using a generalized Pearson system of curves. Then we will work out two cases, the ideal Gaussian state (Sec. II C) and the Gamma state (Sec. II D). In Sec. II E we define the intrinsic entropy function. In the same section we also derive the general expression for the quasi-Gaussian entropy (QGE) and a general differential equation (master equation) which relates the heat capacity and the intrinsic entropy function. In Sec. II F we solve the master equation for the Gaussian and Gamma state which yields the the temperature dependence of the ideal reduced free energy, internal energy and heat capacity. In Secs. II G and II H we describe two approximated solutions which can be applied to systems with more complex statistical states. In Sec. II I we will discuss applications of these approximations to small molecules.

Finally in Sec. III we present the results obtained for three different systems: the classical harmonic solid, for which we could derive an analytical expression for the free energy and the energy distribution, and liquid water and liq-

[a]Department of Chemistry, University of Rome (La Sapienza), p. le A. Moro 5, 00185, Rome, Italy.

 0021-9606/96/104(4)/1560/15/$6.00

uid methanol, where we compared the experimental results with the predictions following from the QGE theory, using both approximations.

II. THEORY

A. Reduced free energy

The Helmholtz free energy of a system with fixed temperature, volume and number of particles, is given by:

$$A=-kT\ \ln\ Q, \tag{2}$$

where Q is the overall partition function. For a system of N identical molecules in the classical limit we can say that

$$Q\cong\frac{1}{N!}Q^{e}Q^{\mathrm{kin}}Q^{\mathrm{pot}}, \tag{3}$$

where Q^e is the electronic partition function, which is a constant for the majority of systems of interest, since the first electronic energy gap is much larger than kT at usual temperature. Q^{kin} is the kinetic energy partition function and Q^{pot} is the potential energy partition function. Q^{kin} and Q^{pot} are partition functions where all the bond lengths and angles and their time derivatives involved in the n_b bond vibrations are considered to be constrained from a classical point of view:

$$Q^{\mathrm{kin}}=\frac{1}{h^R}\int e^{-\beta K}\prod_i^{n_b}\ \delta(\dot{b}_i)d\mathbf{p} \tag{4}$$

$$Q^{\mathrm{pot}}=\sum_{\{l\}}\ \int\ e^{-\beta(\Phi+\psi+E_{\{l\}})}\prod_i^{n_b}\ \delta(b_i-b_i^0)d\mathbf{x}$$

$$=\sum_{\{l\}}\ \int'e^{-\beta(\Phi+\psi+E_{\{l\}})}d\mathbf{x}, \tag{5}$$

where K is the classical expression for the kinetic energy, Φ is the *inter*molecular potential function and ψ is the *intra*molecular potential function (not including bond vibrational potential energy). Further $\delta(b_i-b_i^0)$ and $\delta(\dot{b}_i)$ are Dirac δ-functions representing the bond constraints, where b_i^0 are the equilibrium bond lengths and angles of the quantum oscillators, representing the internal vibrations of the molecules, and $\dot{b}_i$ are the bond constraint time derivatives. $\mathbf{x}$ and $\mathbf{p}$ are the atomic coordinates and momenta, h is Planck's constant, $\beta=1/kT$ and R is the number of the atomic degrees of freedom of the system excluding the vibrational degrees of freedom. $E_{\{l\}}$ is the vibrational energy of the whole system, defined by the set of quantum numbers $\{l\}$ of all quantum oscillators present, and the sum $\Sigma_{\{l\}}$ runs over all possible sets. In general the energy levels of the individual oscillators (and therefore $E_{\{l\}}$) may be a function of the coordinates $\mathbf{x}$, implying a temperature dependence. Only in the ideal gas condition, we can consider the levels of the oscillators to be independent from the coordinates and therefore from the temperature. The prime on the integral sign means an integration with all bond and angle constraints. It must be noted that Q^{kin} in Eq. 4 must be evaluated at one arbitrary configuration, since its Dirac functions $\delta(\dot{b}_i)$ are depending on the coordinates, although the value of the integral is virtually independent of them (since the integration is performed over the whole ensemble of molecules). Therefore the factorization of Q is possible even with constraints.

The free energy of an ideal system at the same temperature and density but with no intermolecular interactions (ideal gas condition) is:

$$A_0=-kT\ \ln\ Q_0, \tag{6}$$

$$Q_0=\frac{1}{N!}Q^{e}Q^{\mathrm{kin}}Q_0^{\mathrm{pot}}, \tag{7}$$

$$Q_0^{\mathrm{pot}}=\sum_{\{l\}}\ \int\ e^{-\beta(\psi+E_{\{l\}0})}\prod_i^{n_b}\ \delta(b_i-b_i^0)d\mathbf{x}$$

$$=\sum_{\{l\}}\ \int'e^{-\beta(\psi+E_{\{l\}0})}d\mathbf{x}. \tag{8}$$

Again, the prime on the integral sign means an integration with all bond and angle constraints. Now the only difference between this ideal state and the real state of the system is in the potential energy partition function. It is convenient to calculate the difference between the real free energy and the ideal one. This difference is called the *reduced free energy*:

$$A^r=A-A_0=-kT\ \ln\left(\frac{Q^{\mathrm{pot}}}{Q_0^{\mathrm{pot}}}\right). \tag{9}$$

In general we can rewrite the (constraint) configurational partition function Q^{pot} for a system with a potential energy function $\nu(\mathbf{x})=\Phi+\psi+E$, with E the total energy of the quantum oscillators, in the following way

$$Q^{\mathrm{pot}}=\sum_{\{l\}}\ \int'e^{-\beta\nu}d\mathbf{x}$$

$$=\left(\sum_{\{l\}}\ \int'e^{-\beta\nu}d\mathbf{x}\right)\frac{\Sigma_{\{l\}}\int'd\mathbf{x}}{\Sigma_{\{l\}}\int'd\mathbf{x}}$$

$$=\frac{\Sigma_{\{l\}}\int'e^{-\beta\nu}d\mathbf{x}\ \Sigma_{\{l\}}\int'd\mathbf{x}}{\Sigma_{\{l\}}\int'e^{\beta\nu}e^{-\beta\nu}d\mathbf{x}}=\frac{\Sigma_{\{l\}}\int'd\mathbf{x}}{\langle e^{\beta\nu}\rangle} \tag{10}$$

and similarly

$$Q_0^{\mathrm{pot}}=\frac{\Sigma_{\{l\}}\int'd\mathbf{x}}{\langle e^{\beta\nu_0}\rangle_0}, \tag{11}$$

where $\langle\cdots\rangle$ and $\langle\cdots\rangle_0$ denote the canonical expectation values in the real and ideal condition, and $\nu_0=\psi+E_0$. It is interesting to note, that the expectation values in Eqs. 10–11 correspond in statistical terms to the moment generating function[8] of the variable ν in β:

$$G_\nu(\beta)=\langle e^{\beta\nu}\rangle=\frac{\Sigma_{\{l\}}\int'd\mathbf{x}}{Q^{\mathrm{pot}}}. \tag{12}$$

We will use this fact later in Sec. II E.

From Eqs. 10 and 11 it follows that

$$A^r=kT\ \ln\ \langle e^{\beta\nu}\rangle-kT\ \ln\ \langle e^{\beta\nu_0}\rangle_0. \tag{13}$$

The two terms on the right hand side are difficult to evaluate, and there is no general way to calculate them. In this article, we will derive for these kind of terms a general analytical expression.

Hence we introduce another reference state, with free energy

$$A_0' = -kT \ln Q_0', \tag{14}$$

$$Q_0' = \frac{1}{N!} Q^e Q^{\mathrm{kin}} \sum_{\{l\}} {\int}' d\mathbf{x}. \tag{15}$$

Note, that this state has the same density and temperature as the real system, but no angle and bond vibrations, and no inter and intramolecular potential, corresponding to an ideal gas of ''rigid'' molecules. The free energy of the real system with respect to this state we call the *ideal reduced free energy* A',

$$A' = A - A_0' = A^r + kT \ln \langle e^{\beta \nu_0} \rangle_0 = kT \ln \langle e^{\beta \nu} \rangle. \tag{16}$$

Equation 16 can be modified to express the energy with respect to the average potential energy U':

$$A' = U' + kT \ln \langle e^{\beta y} \rangle, \tag{17}$$

where

$$U' = \langle \nu \rangle = \langle \Phi + \psi + E \rangle, \tag{18}$$

$$y = \nu - \langle \nu \rangle, \tag{19}$$

and

$$S' = -k \ln \langle e^{\beta y} \rangle \tag{20}$$

is the ideal reduced entropy of the system. This is the difference between the entropy of the real and the new reference system.

It should be noted, that when $\Phi = 0$, Eq. 16 reduces to the second term in Eq. 13. Therefore we can use the expressions we will obtain in the more general case $\Phi \neq 0$ for this term as well. In Sec. II I we will see, that for small molecules A^r can be obtained quite straightforwardly from A'.

B. General distribution and restrictions

It is important to note, that since the energy distribution function $\rho(y)$ is that of a macroscopic system, it must be uninormal and very similar to a Gaussian distribution. We can prove this as follows.

Since the system is macroscopic, we may treat it as a collection of N identical subsystems. Although these subsystems contain a differential amount of matter from a macroscopic point of view, they can be regarded as statistically independent thermodynamic systems since they still contain an ''infinite'' number of molecules. This means that the interaction energy between subsystems is negligible with respect to the internal energy of each subsystem.[9] If therefore every subsystem i has a potential energy ϵ_i and a fluctuation $\Delta \epsilon_i = \epsilon_i - \langle \epsilon_i \rangle$, we can express the potential energy and fluctuation of the total system as

$$\nu = \sum_{i=1}^{N} \epsilon_i, \tag{21}$$

$$\Delta \nu = \nu - \langle \nu \rangle \sum_{i=1}^{N} (\epsilon_i - \langle \epsilon_i \rangle) = \sum_{i=1}^{N} \Delta \epsilon_i. \tag{22}$$

The central limit theorem[10] states, that as $N \to \infty$ (a macroscopic system) $\rho(\Delta \nu) = \rho(y)$ will tend to a Gaussian distribution, whatever is the distribution of $\Delta \epsilon_i$. Hence for a macroscopic system $\rho(y)$ will be uninormal and very similar to a Gaussian distribution. Every macroscopic system which can be described in this way, we call a quasi-Gaussian system, and each type of corresponding energy distribution defines its ''statistical'' state. In Sec. II E we will show that all quasi-Gaussian systems share the same general expression for the ideal reduced entropy.

At very low temperature though, the system must be described in a complete quantum mechanical way. This means, that the energy $\Delta \nu$ and its distribution $\rho(y)$ are discrete instead of continuous, and this discrete distribution need not be uninormal, since at low temperature the degeneracy of energy levels can be small. In general $\rho(y)$ is a product of an increasing function (the degeneracy of levels) and a decreasing function (the Boltzmann factor).

So for systems which are not at too low temperature we can formulate several restrictions on the shape of this near Gaussian energy distribution function:

$[R1]$ The distribution function $\rho(y)$ is defined on the energy interval $[y_0, \infty\rangle$. For physical reasons, y_0 must be a finite value, since any real system must have an overall potential energy minimum. The function therefore cannot be perfectly symmetric. Since for molecular systems at usual temperature, the mode of the energy distribution is very far from the absolute minimum, y_0 can be regarded as a numerical energy minimum, much higher than the real physical minimum. For systems at low density and/or high temperature the mode is so much higher than the numerical minimum that a symmetric curve could be a good description of the real distribution.

$[R2]$ Since we assume the exact probability density (ranging from the absolute minimum to infinity) and at least its first derivative to be continuous, $\rho(y)$ and its derivative $d\rho/dy$ must be continuous and real on the interval $[y_0, \infty\rangle$.

$[R3]$ The derivative $d\rho/dy$ must be zero where ρ equals zero (that is at y_0 and $+\infty$), as follows from $R1$ and $R2$.

$[R4]$ The function $\rho(y)$ is uninormal. It has one extremum at y_m on the interval $[y_0, \infty\rangle$, and this extremum has to be a maximum (the mode). The derivative $d\rho/dy$ therefore must change sign around y_m in the following way:

$$\begin{aligned} \frac{d\rho}{dy} &> 0 \quad (y < y_m), \\ \frac{d\rho}{dy} &< 0 \quad (y > y_m). \end{aligned} \tag{23}$$

$[R5]$ Any acceptable solution of $\rho(y)$ has to result in a finite ideal reduced free energy. Therefore the integral

$$\int_{y_0}^{\infty} e^{\beta y}\rho(y)dy \tag{24}$$

has to converge. (According to Eq. 12 this corresponds to the requirement, that a suitable energy distribution function should have a converging moment generating function at the corresponding β.) So we have the necessary condition that

$$\lim_{y\to\infty} e^{\beta y}\rho(y)=0. \tag{25}$$

It should be mentioned here, that this physical restriction is not valid for a classical ideal harmonic solid (a set of classical oscillators with no limits on the coordinates). We will show in Sec. III, that in such a case the system still can be described as a special case of the theory.

The most general differential equation, describing such a uninormal function, is

$$\frac{d\rho}{dy}=-(y-y_m)\rho\frac{P^m(y)}{G^n(y)} \tag{26}$$

with $P^m(y)$ and $G^n(y)$ some arbitrary polynomials in y of order m and n:

$$P^m(y)=\sum_{i=0}^{m} a_i y^i, \tag{27}$$

$$G^n(y)=\sum_{j=0}^{n} b_j y^j. \tag{28}$$

Eq. 26 is just a generalization of the normal Pearson system of frequency curves, developed by Karl Pearson between 1890 and 1900.[11–13] It is based on the differential equation

$$\frac{d\rho}{dy}=-\frac{y-y_m}{b_0+b_1y+b_2y^2}\rho \tag{29}$$

corresponding to $m=0$, $n=2$ in our notation. The choice $b_2=0$ or $b_2=b_1=0$ is equal to $n=1$ or $n=0$. The resulting family includes some well-known types, like the normal, Gamma, Beta and t-distribution.[14,15] Further generalizations of Eq. 29 are described by Ord,[14] for the case $n>2$. The use of a Padé expansion as in Eq. 26 was first described by Dunning and Hanson.[16] We will use the generalized Pearson equation to classify different uninormal distributions, close to a Gaussian, which satisfy the physical and mathematical restrictions.

Following Pearson's work we can moreover link the parameters y_m, a_i, and b_j to the (central) moments of the potential energy M_n, defined as

$$M_n=\langle(\nu-\langle\nu\rangle)^n\rangle=\langle y^n\rangle. \tag{30}$$

If we have n_p *independent* parameters, we have to solve the following set of $k_{\max}$ equations

$$\int_{y_0}^{\infty} y^k G^n(y)\frac{d\rho}{dy}dy=-\int_{y_0}^{\infty} y^k(y-y_m)P^m(y)\rho dy \qquad k=0\ldots k_{\max}-1 \tag{31}$$

with

$$k_{\max}=n_e\geqslant n_p, \tag{32}$$

where n_e is the number of equations needed to solve the parameters. n_e can be larger than n_p because the first n_p equations not necessarily are independent. Integrating by parts we obtain

$$[y^kG^n(y)\rho]_{y_0}^{\infty}-\int_{y_0}^{\infty}\rho\frac{d}{dy}\{y^kG^n(y)\}dy = -\int_{y_0}^{\infty} y^k(y-y_m)P^m(y)\rho dy. \tag{33}$$

Since all moments of finite order exist, as it follows from the fact that we can express every moment M_{n+1} as a sum of products of $\langle\nu\rangle$ and its first n derivatives in temperature, we can conclude that $\lim_{y\to\infty}y^kG^n(y)\rho=0$. Hence, since ρ is zero at y_0

$$\int_{y_0}^{\infty}\rho\frac{d}{dy}\{y^kG^n(y)\}dy=\int_{y_0}^{\infty} y^k(y-y_m)P^m(y)\rho dy. \tag{34}$$

Using the definitions of $G^n(y)$ and $P^m(y)$, we get

$$\int_{y_0}^{\infty}\sum_{j=0}^{n}(j+k)b_jy^{j+k-1}\rho dy = \int_{y_0}^{\infty}\sum_{i=0}^{m} a_i\{y^{i+k+1}-y_my^{i+k}\}\rho dy, \tag{35}$$

which results in the following set of relations between the parameters and moments

$$\sum_{j=0}^{n}(j+k)b_jM_{j+k-1}=\sum_{i=0}^{m} a_i\{M_{i+k+1}-y_mM_{i+k}\} \qquad k=0\ldots k_{\max}-1 \tag{36}$$

with $M_0=1$, and $M_1=0$.

Later on we will see, that $m+1\geqslant n$ (Eqs. 41 and 42); therefore the highest moment that is required is $M_{k_{\max}+m}$. From the same equations as Eq. 36 with $k>k_{\max}-1$, we get the moments from $M_{k_{\max}+m+1}$ on, expressed in terms of the lower ones.

We will now investigate the effect of the restrictions $R1$ to $R5$ on the possible solutions of the differential equation of $\rho(y)$, Eq. 26.

Restriction $R2$ (continuity) demands that $G^n(y)$ has no real roots on the interval $[y_0,\infty\rangle$. $G^n(y)$ therefore can have real roots only at y_0 or some value $\bar{y_i}$ smaller than y_0. Since $\rho(y)$ and its derivative must be real, also pairs of conjugated complex roots at $\eta_i\pm i\hat{\eta}_i$ are possible. So $G^n(y)$ can be built up of factors like $(y-y_0)^{n'}$, $(y-\bar{y_i})^{n''}$, and $\Pi_i^{n'''}((y-\eta_i)^2+\hat{\eta}_i^2)$ with $n'+n''+n'''/2=n$. None of these factors can change sign in the interval $[y_0,\infty\rangle$.

Similarly, restriction $R4$ demands that $P^m(y)$ has no real roots on the interval $[y_0,\infty\rangle$ other than y_m. $P^m(y)$ can have other real roots smaller than or equal to y_0, or has pairs of conjugated complex roots at $\epsilon_i\pm i\hat{\epsilon}_i$. Therefore $P^m(y)$ can

consist of factors like $(y-y_m)^{m'}$, $(y-y_0)^{m''}$, $(y-\bar{y}_i)^{m'''}$, and $\Pi_i^{m'v}((y-\epsilon_i)^2+\hat{\epsilon}_i^2)$. Of these factors, only $(y-y_m)^{m'}$ can change sign in the interval. Since $\rho(y)$ must have a maximum, restriction $R4$ demands that $P^m(y)/G^n(y)>0$, so m' = even.

Restriction $R5$ (convergence) requires that Eq. 25 is fulfilled. We can use the formal solution of Eq. 26 for $\rho(y)$:

$$\rho(y)=\rho_0\exp\left\{-\int (y-y_m)\,\frac{P^m(y)}{G^n(y)}dy\right\}. \tag{37}$$

If we combine this with Eq. 25 we obtain

$$\lim_{y\to\infty}\ \rho_0\ \exp\left\{\beta y-\int (y-y_m)\frac{P^m(y)}{G^n(y)}dy\right\}=0, \tag{38}$$

which is equivalent to

$$\lim_{y\to\infty}\int (y-y_m)\,\frac{P^m(y)}{G^n(y)}dy-\beta y=\infty,$$

$$\lim_{y\to\infty}\beta y\times\lim_{y\to\infty}\left(\frac{\int (y-y_m)\frac{P^m(y)}{G^n(y)}dy}{\beta y}-1\right)=\infty,$$

$$\lim_{y\to\infty}\frac{\int (y-y_m)\frac{P^m(y)}{G^n(y)}dy}{\beta y}>1. \tag{39}$$

Using l'Hôpital's rule we finally obtain

$$\lim_{y\to\infty}\frac{G^n(y)}{(y-y_m)P^m(y)}<\beta^{-1}=kT. \tag{40}$$

This inequality is always fulfilled if

$$m+1>n. \tag{41}$$

It is never fulfilled when $m+1<n$. In the case that $m+1=n$, we use the definitions of the polynomials $P^m(y)$ and $G^n(y)$ (Eqs. 27 and 28) and find that the free energy can converge only if

$$\frac{b_n}{a_m}<kT\quad m=n-1. \tag{42}$$

We can order the possible solutions of Eq. 26 according to their complexity. This is defined by the number of parameters $(=m+n+2)$ and the number of required moments $(=k_{\max}+m)$.

The simplest solution of Eq. 26 $\{m=0,\ n=0\}$ is a Gaussian distribution (defining the ideal Gaussian state) and the first completely acceptable solution $\{m=0,\ n=1\}$ is a Gamma distribution, defining the Gamma state. In this way, a state is defined by the type of potential energy distribution, because it describes all the equilibrium properties of the system. In the following sections we will focus on these two solutions.

C. Ideal Gaussian state

From the central limit theorem we saw that $\rho(y)$ tends to a Gaussian distribution as the system becomes more and more macroscopic. Therefore we can consider as ideal reference a system where $\rho(y)$ is exactly a Gaussian. We call this the ideal *Gaussian state*.

This case corresponds to $\{m=0,\ n=0\}$. It is the simplest solution, with two parameters (since we can put a_0 to 1) and two moments, with the following differential equation for $\rho(y)$:

$$\frac{d\rho}{dy}=-\frac{(y-y_m)}{b_0}\rho. \tag{43}$$

Eq. 36 yields the following set of relations between parameters and moments:

$$\begin{cases}0=-y_m\\ b_0=M_2\end{cases}. \tag{44}$$

The differential equation thus becomes

$$\frac{d\rho}{dy}=-\frac{y}{M_2}\rho, \tag{45}$$

which has the trivial solution

$$\rho(y)=\rho_0 e^{-y^2/2M_2} \tag{46}$$

with $\rho_0=1/\sqrt{2\pi M_2}$. It should be noted, that a real Gaussian state is not possible since it would imply a perfectly symmetric distribution with no energy limits. This is physically impossible because every system must have an overall energy minimum (restriction $R1$). Under certain conditions though, it might be a good numerical approximation.

The ideal reduced Helmholtz free energy is calculated as follows, according to Eq. 17

$$A'=U'+kT\ \ln\left(\frac{1}{\sqrt{2\pi M_2}}\int_{-\infty}^{\infty}e^{\beta y}e^{-y^2/2M_2}dy\right). \tag{47}$$

Using the substitution $z=1/\sqrt{2\pi M_2}(y-\beta M_2)$, we obtain

$$A'=U'+\frac{M_2}{2kT}, \tag{48}$$

$$S'=-\frac{M_2}{2kT^2}. \tag{49}$$

Eq. 49 shows that in an ideal Gaussian state the reduced entropy of the system is completely defined by T and the second central moment of the potential energy.

It is clear then, that in such a condition A' is a function only of T and the first two potential energy moments. We can link these moments M_n to the derivatives of $\langle\nu\rangle$ in T:

$$\left(\frac{\partial\langle\nu\rangle}{\partial\beta}\right)_V=-kT^2\left(\frac{\partial\langle\nu\rangle}{\partial T}\right)_V=-kT^2C_V' \tag{50}$$

$$\left(\frac{\partial\langle\nu\rangle}{\partial\beta}\right)_V=\frac{\partial}{\partial\beta}\left\{\frac{\Sigma\int'\nu e^{-\beta\nu}d\mathbf{x}}{\Sigma\int' e^{-\beta\nu}d\mathbf{x}}\right\}=\langle\nu\rangle^2-\langle\nu^2\rangle=-M_2. \tag{51}$$

Here again the prime on the integration sign means integrating with all bond and angle constraints, and the summation runs over all possible sets of quantum vibrational states. So

$$M_2 = kT^2 C_V', \tag{52}$$

$$A' = U' + \frac{TC_V'}{2}, \tag{53}$$

$$S' = -\frac{C_V'}{2}. \tag{54}$$

Here C_V' is the *ideal reduced* heat capacity of the total system at fixed volume, excluding the kinetic energy part.

D. Gamma state

The first physically acceptable solution of ρ just beyond the Gaussian distribution (Eq. 46) is given by $\{m=0, n=1\}$. In this case we have 3 independent parameters (putting a_0 to 1 again) and 3 potential energy moments. The differential equation of ρ is

$$\frac{d\rho}{dy} = -\frac{(y-y_m)}{b_0+b_1 y}\rho = -\frac{(y-y_m)}{b_1(y-y_0)}\rho, \tag{55}$$

where $y_0 = -b_0/b_1$. Using Eq. 36, the three parameters y_m, b_0, and b_1 are linked to the energy moments in the following way:

$$\begin{cases} b_1 = -y_m \\ b_0 = M_2 \\ 3b_1 M_2 = M_3 - y_m M_2 \ . \end{cases} \tag{56}$$

Hence we have

$$b_0 = M_2, \tag{57}$$

$$b_1 = -y_m = \frac{M_3}{2M_2}. \tag{58}$$

From Eq. 36 for $k > k_{\max} - 1$ we obtain the recursive relations that link the moments higher than M_3 to M_2 and M_3. It turns out that we can express M_n as

$$M_n = (n-1)\left\{ M_2 M_{n-2} + \frac{M_3}{2M_2} M_{n-1} \right\}. \tag{59}$$

Note the similarity between this expression and the one obtained for the Gaussian distribution,

$$\begin{aligned} &M_n = (n-1) M_2 M_{n-2}, \\ &M_{odd} = 0. \end{aligned} \tag{60}$$

The distribution function follows from

$$\frac{d\rho}{dy} = -\frac{y+b_1}{b_0+b_1 y}\rho \tag{61}$$

and is only defined in $[y_0, \infty\rangle$. For convenience, we define $z = b_0 + b_1 y$, $z > 0$, and obtain

$$b_1 \frac{d\rho}{dz} = \left\{ -\frac{1}{b_1} + \frac{b_0 - b_1^2}{b_1 z} \right\}\rho. \tag{62}$$

By integrating this equation, we find

$$\rho(z) = \rho_0 z^{b_0/b_1^2 - 1} e^{-z/b_1^2}, \tag{63}$$

where ρ_0 is the normalization constant, defined as

$$\begin{aligned} \frac{1}{\rho_0} &= \int_0^\infty z^{b_0/b_1^2 - 1} e^{-z/b_1^2} dz \\ &= (b_1^2)^{b_0/b_1^2} \int_0^\infty x^{b_0/b_1^2 - 1} e^{-x} dx = (b_1^2)^{b_0/b_1^2} \Gamma\left(\frac{b_0}{b_1^2}\right), \end{aligned} \tag{64}$$

where we used the substitution $x = z/b_1^2$. In the last step we used the definition of the Gamma function.[17] Now the distribution function ρ can be expressed as either

$$\rho(z) = \frac{\left(\frac{1}{b_1^2}\right)^{b_0/b_1^2}}{\Gamma\left(\frac{b_0}{b_1^2}\right)} z^{b_0/b_1^2 - 1} e^{-z/b_1^2} \tag{65}$$

or

$$\rho(y) = \frac{b_1 \left(\frac{1}{b_1^2}\right)^{b_0/b_1^2}}{\Gamma\left(\frac{b_0}{b_1^2}\right)} (b_0 + b_1 y)^{b_0/b_1^2 - 1} e^{-(b_0 + b_1 y)/b_1^2}, \tag{66}$$

where we used $\rho(z)dz = \rho(y)dy$. Eqs. 65 or 66 represent the Gamma distribution,[8] or Pearson's type III curve, which has the general form

$$\rho(\xi) = \frac{\theta^{\,a}}{\Gamma(a)} \xi^{a-1} e^{-\theta\xi}, \quad \xi > 0. \tag{67}$$

In this case, $\xi = z = b_0 + b_1 y$, $a = b_0/b_1^2$ and $\theta = 1/b_1^2$.

The ideal reduced Helmholtz free energy can be calculated from Eq. 17; therefore we need to evaluate

$$\begin{aligned} \langle e^{\beta y} \rangle &= \int_{y_0}^\infty e^{\beta y} \rho(y) dy \\ &= \int_0^\infty e^{\beta(z-b_0)/b_1} \rho(z) dz \\ &= e^{-\beta b_0/b_1} \frac{\left(\frac{1}{b_1^2}\right)^{b_0/b_1^2}}{\Gamma\left(\frac{b_0}{b_1^2}\right)} \int_0^\infty z^{b_0/b_1^2 - 1} e^{-(1/b_1^2 - \beta/b_1)z} dz \\ &= e^{-\beta b_0/b_1} \frac{(1-b_1\beta)^{-b_0/b_1^2}}{\Gamma\left(\frac{b_0}{b_1^2}\right)} \int_0^\infty x^{b_0/b_1^2 - 1} e^{-x} dx \\ &= e^{-\beta b_0/b_1} (1-b_1\beta)^{-b_0/b_1^2}, \end{aligned} \tag{68}$$

where we used $x = (1-b_1\beta)z/b_1^2$. Hence the free energy is

$$A' = U' + kT \ln \langle e^{\beta y} \rangle$$
$$= U' + kT \left\{ -\beta \frac{b_0}{b_1} - \frac{b_0}{b_1^2} \ln (1 - b_1 \beta) \right\}$$
$$= U' - \frac{b_0}{b_1} - \frac{b_0}{b_1^2} kT \ln \left(1 - \frac{b_1}{kT} \right). \tag{69}$$

Since $m+1=n$, we have the additional restriction that $b_n/a_m = b_1 < kT$ (see Eq. 42). This also follows from Eq. 69, since the argument of the logarithm cannot be negative or zero. As already mentioned in Sec. II B a classical ideal harmonic solid has exactly the energy distribution of the Gamma state (Eq. 65), but with $b_1 = kT$, resulting in a diverging ideal reduced free energy, see Sec. III.

Remembering the expressions of b_0 and b_1 (Eqs. 57 and 58), it is clear that in the Gamma state the free energy is given by U', M_2 and M_3. Just as we did previously, we can express M_2 and M_3 as temperature derivatives of $\langle \nu \rangle$. Using Eqs. 50 and 51 we see that

$$\left(\frac{\partial^2 \langle \nu \rangle}{\partial \beta^2} \right)_V = \frac{\partial}{\partial \beta} \{ -kT^2 C_V' \} = (kT^2)^2 \left(\frac{\partial C_V'}{\partial T} \right)_V + 2k^2 T^3 C_V' \tag{70}$$

and

$$\left(\frac{\partial^2 \langle \nu \rangle}{\partial \beta^2} \right)_V = \frac{\partial}{\partial \beta} \left\{ \left(\frac{\Sigma \int' \nu e^{-\beta \nu} d\mathbf{x}}{\Sigma \int' e^{-\beta \nu} d\mathbf{x}} \right)^2 - \frac{\Sigma \int' \nu^2 e^{-\beta \nu} d\mathbf{x}}{\Sigma \int' e^{-\beta \nu} d\mathbf{x}} \right\}$$
$$= \langle \nu^3 \rangle - 3 \langle \nu^2 \rangle \langle \nu \rangle + 2 \langle \nu \rangle^3$$
$$= M_3 \tag{71}$$

so

$$M_3 = (kT^2)^2 \left(\frac{\partial C_V'}{\partial T} \right)_V + 2k^2 T^3 C_V'. \tag{72}$$

In the Gamma state therefore, the ideal reduced free energy is given by the average potential energy and its first two derivatives in temperature.

Now we can use the expressions for M_2 (Eq. 52) and M_3 in the definition of b_0 and b_1. It follows that

$$A' = U' - TC_V' \left\{ \frac{kT}{b_1} + \left(\frac{kT}{b_1} \right)^2 \ln \left(1 - \frac{b_1}{kT} \right) \right\} \tag{73}$$
$$= U' - TC_V' \alpha(\delta) \tag{74}$$

with

$$\delta(T) = \frac{b_1}{kT} = 1 + \frac{T}{2C_V'} \left(\frac{\partial C_V'}{\partial T} \right)_V, \tag{75}$$

$$\alpha(\delta) = \frac{1}{\delta} + \left(\frac{1}{\delta} \right)^2 \ln (1 - \delta). \tag{76}$$

Since $b_1 < kT$, it follows that $0 < \delta < 1$ and from Eq. 75 it is clear that $(\partial C_V'/\partial T)_V < 0$, implying that $C_V' \to 0$ as $T \to \infty$.

The function $\alpha(T)$ plays a very important role in the theory. Among others it defines the ideal reduced entropy as

$$S' = C_V' \alpha(T). \tag{77}$$

Inspecting Eqs. 75 and 76 we see, that α is actually a function of $T(\partial C_V'/\partial T)_V/2C_V'$. It is very important to note that $\alpha(T)$ is a dimensionless and intensive quantity, independent of the size of the system. Comparing Eqs. 48 and 74 we see that the ideal Gaussian state implies that $\alpha(T) = -1/2$. Such a condition can be regarded as a limit of Eq. 76 when $b_1 \to 0$. The parameter b_1 is therefore a measure of the asymmetry of the energy distribution.

E. The intrinsic entropy function

In the previous sections we studied the first two possible statistical states of a quasi-Gaussian system. Although it is possible to study higher order solutions of the generalized Pearson equation, such solutions will be very complicated, involving higher order moments (and derivatives of $\langle \nu \rangle$ in T). It turns out that solutions just beyond the Gamma distribution involve at least M_5.

In this section we will show that for every quasi-Gaussian system α can be expressed in terms of the moment generating function of the subsystems. We will also derive a general differential equation (thermodynamic master equation) which links $\alpha(T)$ to $C_V(T)$.

In Sec. II B we defined a quasi-Gaussian system as a collection of N identical independent differential thermodynamic subsystems (elementary systems), with $N \to \infty$. We can define the moment generating function of the energy fluctuation $\Delta\epsilon$ for each elementary system as

$$g_{\Delta\epsilon}(\beta) = 1 + \mu_2 \frac{\beta^2}{2} + \mu_3 \frac{\beta^3}{3!} + \cdots + \mu_k \frac{\beta^k}{k!} + \cdots \tag{78}$$

with $\mu_k = \langle (\Delta\epsilon)^k \rangle$ the kth potential energy moment of the elementary system.

For the whole system the moment generating function (being the product of the moment generating functions of the independent subsystems[18]) and the ideal reduced entropy (see Eq. 20) are

$$G_{\Delta\nu}(\beta) = \lim_{N \to \infty} (g_{\Delta\epsilon}(\beta))^N,$$
$$S' = -Nk \ln (g_{\Delta\epsilon}(\beta)). \tag{79}$$

Since

$$C_V' = \frac{\partial \langle \nu \rangle}{\partial T} = N \frac{\partial \langle \epsilon \rangle}{\partial T}, \tag{80}$$

it follows that

$$\frac{S'}{C_V'} = \frac{-k \ln (g_{\Delta\epsilon}(\beta))}{\partial \langle \epsilon \rangle / \partial T}$$
$$= -\frac{\ln \left(1 + \mu_2 \frac{\beta^2}{2} + \mu_3 \frac{\beta^3}{3!} + \ldots \right)}{\mu_2 \beta^2} = \alpha(T). \tag{81}$$

It is evident that α is a dimensionless and intensive property, since it is completely defined by the elementary system. In general α is a function of the temperature, density and com-

position of the system. We can call this quantity the *intrinsic entropy function*. It is interesting to note that $\alpha<0$ (since $S'<0$ and $C_V'>0$) and in the limit of infinite temperature we have

$$\lim_{T\to\infty}\alpha(T)=\lim_{\beta\to 0}-\frac{\ln\left(1+\mu_2\frac{\beta^2}{2}+\mu_3\frac{\beta^3}{3!}+\ \ldots\right)}{\mu_2\beta^2}=-\frac{1}{2} \tag{82}$$

as it follows from the fact that every μ_k converges to a finite value as $T\to\infty$ (being equal to $\langle(\Delta\epsilon)^k\rangle_0$).

From Eq. 81 we can also define the ideal reduced free energy as

$$A'=U'+kT\ \ln\ G_{\Delta\nu}(\beta)=U'-TC_V'\alpha(T), \tag{83}$$

where $C_V'\alpha$ is the general expression for the quasi-Gaussian entropy. $\alpha(T)$ is given by $-1/2$ for a Gaussian state and by Eq. 76 for the Gamma state. Every type of distribution will have its own functional form of α depending on a limited number of moments. This number increases with increasing order of the solution of the generalized Pearson equation. It is interesting to note that solutions of lower order can be regarded as limits of higher order solutions. With increasing temperature, the system can be described by a lower order solution and in the limit of infinite temperature every system will tend to an ideal Gaussian state.

Using Eq. 83, we can derive a new differential equation, since we know from thermodynamics that

$$\left(\frac{\partial A'}{\partial T}\right)_V=-S'=-C_V'\alpha(T). \tag{84}$$

Differentiating Eq. 83 we obtain on the other hand

$$\left(\frac{\partial A'}{\partial T}\right)_V=C_V'-C_V'\alpha(T)-T\alpha(T)\left(\frac{\partial C_V'}{\partial T}\right)_V - TC_V'\left(\frac{\partial\alpha(T)}{\partial T}\right)_V \tag{85}$$

and hence we arrive at the *thermodynamic master equation*

$$T\alpha(T)\left(\frac{\partial C_V'}{\partial T}\right)_V=C_V'-TC_V'\left(\frac{\partial\alpha(T)}{\partial T}\right)_V. \tag{86}$$

It is a remarkable fact that if the exact statistical state of a quasi-Gaussian system is known at least at one temperature, Eq. 86 will provide the complete behavior of C_V' and hence of all other thermodynamical properties at every possible temperature at fixed volume, if of course there are no singular points, where $(\partial C_V'/\partial T)_V=\pm\infty$ or not continuous. It should be mentioned also that the derivations of this section are valid even when the probability distribution function of the potential energy is not continuous and/or uninormal.

F. Exact solutions of the thermodynamic master equation

In this section we will solve the master equation, Eq. 86, for the two cases we discussed before in the previous sections. For a Gaussian state, since $\alpha=-1/2$ and hence $(\partial\alpha/\partial T)_V=0$ the thermodynamic master equation is simply

$$\frac{dC_V'}{C_V'}=-\frac{2}{T}dT \tag{87}$$

with solution of $C_V'(T)$

$$C_V'(T)=C_V'(T_0)\left(\frac{T_0}{T}\right)^2. \tag{88}$$

The potential energy is

$$U'(T)=U'(T_0)+\int_{T_0}^{T}C_V'(T)dT =U'(T_0)+T_0C_V'(T_0)\frac{(T-T_0)}{T} \tag{89}$$

and free energy

$$A'(T)=U'(T)-TC_V'(T)\alpha(T) =U'(T_0)+T_0C_V'(T_0)\left[1-\frac{T_0}{2T}\right]. \tag{90}$$

In the case of the Gamma state, we can rewrite the master equation in terms of δ (Eq. 75) as

$$T\left(\frac{\partial\alpha}{\partial T}\right)=2(1-\delta)\alpha+1, \tag{91}$$

where $(\partial\alpha/\partial T)=(\partial\alpha/\partial\delta)(\partial\delta/\partial T)$. Since from Eq. 76

$$\left(\frac{\partial\alpha}{\partial\delta}\right)=-\frac{2\alpha(1-\delta)+1}{\delta(1-\delta)} \tag{92}$$

we get $(\partial\delta/\partial T)=-\delta(1-\delta)/T$, so

$$\int_{T_0}^{T}\frac{dT}{T}=-\int_{T_0}^{T}\frac{d\delta}{\delta(1-\delta)} \tag{93}$$

with solution

$$\delta(T)=\frac{T_0\delta_0}{T(1-\delta_0)+T_0\delta_0}, \tag{94}$$

where $\delta_0=\delta(T_0)$. From Eq. 75 it follows:

$$C_V'(T)=C_V'(T_0)\left(\frac{T_0}{T(1-\delta_0)+T_0\delta_0}\right)^2, \tag{95}$$

$$U'(T)=U'(T_0)+T_0C_V'(T_0)\ \frac{(T-T_0)}{T(1-\delta_0)+T_0\delta_0}, \tag{96}$$

$$A'(T)=U'(T_0)-\frac{T_0C_V'(T_0)}{\delta_0}-\frac{TC_V'(T_0)}{\delta_0^2} \times\ln\left(\frac{T(1-\delta_0)}{T(1-\delta_0)+T_0\delta_0}\right). \tag{97}$$

From the solutions it is clear that in the Gaussian state the knowledge of U' and C_V' at one temperature is required to know the complete behavior of the system. In the Gamma state we need to know $(\partial C_V'/\partial T)$ at one temperature as well.

The Gaussian state and its solution, as we mentioned before, can be regarded as an ideal solution. The Gamma state solution on the contrary, although in general it will not be an exact description, can be considered as an acceptable approximation for a large variety of systems. In the next section we will describe an approximated solution of the thermodynamic master equation, based on the assumption that the state of a real system is very close to a Gamma state.

G. Effective Gamma approximation

In the previous section we derived the solution for an exact Gamma state. In this section we describe an approximation for real systems which can not be described by the exact Gamma solution, but which can be considered to be in a perturbed Gamma state.

In general the heat capacity of the real system can be written as

$$C'_V(T)=C'_{V\Gamma}(T)+\Delta(T), \tag{98}$$

where $C'_{V\Gamma}$ is the heat capacity of a possible Gamma state and $\Delta(T)$ the corresponding error. Our aim is to find a specific ''effective'' Gamma state which will reproduce the entropy of the real system as close as possible. If all the information on the system is confined to one temperature, one simple choice to define uniquely the Gamma state is to equate $S'_\Gamma(T_0)$ with $S'(T_0)$ and $C'_{V\Gamma}(T_0)$ with $C'_V(T_0)$. This effective Gamma state is thus defined as having the same entropy and heat capacity as the real system at the reference temperature T_0. Therefore we can use Eqs. 94–97 replacing δ and δ_0 by the effective δ^* and δ_0^*, where δ_0^* follows from $S'(T_0)/C'_V(T_0)=\alpha(T_0)=(\delta_0^*)^{-1}+(\delta_0^*)^{-2}\ln(1-\delta_0^*)$. This effective Gamma state could provide a good approximation for the entropy (especially for $T>T_0$) and therefore the free energy, since $A'(T)-A'(T_0)=-\int_{T_0}^{T}S'(T)dT \cong -\int_{T_0}^{T}S'_\Gamma(T)dT$. It should be noted that the average energy will not be reproduced equally well, since $\int_{T_0}^{T}\Delta(T)dT$ is in general not negligible.

H. Constant alpha approximation

When the statistical state of a quasi-Gaussian system is known (and hence the functional form of $\alpha(T,C'_V,(\partial C'_V/\partial T),\ldots)$) the thermodynamic master equation (Eq. 86) can be solved exactly, as we did for the Gaussian and Gamma state. In the previous section we showed that an approximated solution can be obtained if the system is close to a Gamma state (perturbed Gamma state). In this section we will derive an approximated (local) solution of the master equation in some temperature interval based on a temperature expansion of the intrinsic entropy function. This approximation does not require any knowledge or assumption on the statistical state of the system.

We can expand $\alpha(T)$ in T around a reference temperature T_0

$$\alpha(T)=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\left.\frac{\partial^n\alpha(T)}{\partial T^n}\right|_{T_0}\right)_V(T-T_0)^n \tag{99}$$

and then use this to integrate Eq. 86. With Eq. 99 we can produce a set of approximated analytical solutions, valid for an increasing temperature range. We started testing the zero order solution, where $\alpha(T)\cong\alpha(T_0)$ and $(\partial\alpha(T)/\partial T)_V\cong 0$. We obtained already with this simple solution very accurate results for a large temperature range for two real liquids.

Within the zero order approximation, we can rewrite Eq. 86 as

$$\left(\frac{\partial C'_V}{\partial T}\right)_V=\frac{C'_V}{\alpha(T_0)T}, \tag{100}$$

which has the solution

$$C'_V(T)=C'_V(T_0)\left(\frac{T}{T_0}\right)^{\lambda}, \tag{101}$$

$$\lambda=\frac{1}{\alpha(T_0)}. \tag{102}$$

We can use Eq. 101 to calculate U' as a function of T:

$$\begin{aligned}U'(T)&=U'(T_0)+\int_{T_0}^{T}C'_V(T)dT\\&=U'(T_0)+C'_V(T_0)\int_{T_0}^{T}\left(\frac{T}{T_0}\right)^{\lambda}dT\\&=U'(T_0)+\frac{C'_V(T)T}{\lambda+1}-\frac{C'_V(T_0)T_0}{\lambda+1}.\end{aligned} \tag{103}$$

Eq. 103 can also be used to calculate A' from Eq. 74

$$\begin{aligned}A'&=U'(T_0)+\frac{C'_V(T)T}{\lambda+1}-\frac{C'_V(T_0)T_0}{\lambda+1}-\frac{C'_V(T)T}{\lambda}\\&=U'(T_0)-\frac{C'_V(T_0)}{\lambda+1}\left\{\alpha(T_0)T\left(\frac{T}{T_0}\right)^{\lambda}+T_0\right\}.\end{aligned} \tag{104}$$

Eqs. 101, 103 and 104 give C'_V, U' and A' as a function of temperature in the zero order approximation of Eq. 86.

I. Application to small molecules

In this section we will apply the results of the quasi-Gaussian Entropy theory to small molecules, like water and methanol. We can rewrite Eq. 13, expressing the first two terms on the right hand side according to Eq. 83 as:

$$A^r=U'-U'_0-T\{C'_V\alpha(T)-C'_{V0}\alpha_0(T)\}, \tag{105}$$

where

$$U'_0=\langle\nu_0\rangle_0, \tag{106}$$

$$C'_{V0}=\left(\frac{\partial U'_0}{\partial T}\right)_V. \tag{107}$$

For an ideal gas, $Q_0^{\rm pot}$ can be factorized as

$$Q_0^{\rm pot}=Q_0^{v}\int' e^{\beta\psi}d\mathbf{x}. \tag{108}$$

Now Q_0^v is the product of all molecular vibrational partition functions q_0^v, where we can approximate q_0^v in terms of quantum oscillators[5]

$$Q_0^v=(q_0^v)^N=\left(\prod_i^m q_{0i}^v\right)^N,$$
$$q_{0i}^v=\frac{e^{-\beta e_{0i}^0}}{1-e^{-2\beta e_{0i}^0}} \tag{109}$$

with e_{0i}^0 the zero-point energy of the ith harmonic mode, and m the number of modes per molecule. The possibility to factorize $Q_0^{\rm pot}$ implies, that e_{0i}^0 is independent of $\mathbf{x}$ and therefore of T. As explained before, the levels in the real system may be temperature dependent. For water for example, there is a shift of one of the three modes of about 100 cm^{-1} in 400 K.[19]

It is possible to calculate Q_0^v using spectroscopical data, obtaining that for most molecules up to at least 600 K virtually no other vibrational states other than the ground state are populated, since $e_{i\,0}^0 \gg kT$ (Ref. 20) (except for backbone vibrations in large molecules[21]). For water[22] and methanol,[23] this is fulfilled even up to 1000 K. Therefore, the average vibrational energy is just the ground state energy, and the vibrational contribution to the heat capacity is virtually zero.

Moreover, for small molecules, like water and methanol (for the latter considering the dihedral angle to be freely rotating), we have $\psi=0$, and therefore

$$U_0'=\langle \nu_0\rangle_0=N\sum_i^m e_{0i}^0=E_0^0, \tag{110}$$

$$C_{V0}'=0, \tag{111}$$

and so

$$U'-U_0'=U'-E_0^0=U^r, \tag{112}$$

$$C_V'=C_V^r, \tag{113}$$

$$A^r=U^r-TC_V^r\alpha(T), \tag{114}$$

where U^r is the vaporization energy and $C_V^r=(\partial U^r/\partial T)_V$. Hence for pure systems, consisting of small molecules we can express the equations of the effective Gamma approximation as

$$\delta^*(T)=\frac{T_0\delta_0^*}{T(1-\delta_0^*)+T_0\delta_0^*}, \tag{115}$$

$$c_V^r(T)=c_V^r(T_0)\left(\frac{T_0}{T(1-\delta_0^*)+T_0\delta_0^*}\right)^2, \tag{116}$$

$$U_m^r(T)=U_m^r(T_0)+T_0c_V^r(T_0)\,\frac{(T-T_0)}{T(1-\delta_0^*)+T_0\delta_0^*}, \tag{117}$$

$$A_m^r(T)=U_m^r(T_0)-\frac{T_0c_V^r(T_0)}{\delta_0^*}-\frac{Tc_V^r(T_0)}{(\delta_0^*)^2}\times\ln\left(\frac{T(1-\delta_0^*)}{T(1-\delta_0^*)+T_0\delta_0^*}\right), \tag{118}$$

where δ_0^* follows from $\alpha(T_0)=(\delta_0^*)^{-1}+(\delta_0^*)^{-2}\ln(1-\delta_0^*)$, and modify the constant alpha equations (Eqs. 101–104) to obtain

$$c_V^r(T)=c_V^r(T_0)\left(\frac{T}{T_0}\right)^\lambda, \tag{119}$$

$$U_m^r(T)=U_m^r(T_0)+\frac{c_V^r(T)T}{\lambda+1}-\frac{c_V^r(T_0)T_0}{\lambda+1}, \tag{120}$$

$$A_m^r=U_m^r(T_0)-\frac{c_V^r(T_0)}{\lambda+1}\left\{\alpha(T_0)T\left(\frac{T}{T_0}\right)^\lambda+T_0\right\}, \tag{121}$$

$$\lambda=\frac{1}{\alpha(T_0)} \tag{122}$$

with c_V^r, U_m^r and A_m^r the molar reduced heat capacity, potential energy and free energy.

III. RESULTS

In this section we will present the results of the quasi-Gaussian Entropy theory (QGE) on three systems. First we will discuss the harmonic solid (a generalization of the classical Einstein solid). Then we present data on liquid water and methanol, using both the effective Gamma and constant alpha approximations (for these liquids an exact Gamma solution is not accurate enough).

The classical Einstein solid is a collection of one-dimensional identical oscillators, which are coupled to a thermal bath.[24] Here we will investigate a collection of non-identical classical harmonic oscillators. The potential energy of the whole system, Φ, is a simple square potential,

$$\Phi=\frac{1}{2}\sum_{i=1}^{\nu}\kappa_i x_i^2, \tag{123}$$

where ν is the number of oscillators, κ_i is the force constant of the ith oscillator, and x_i the corresponding coordinate. The configurational part of the free energy is therefore

$$\begin{aligned}A&=-kT\ln Q\\&=-kT\ln\int e^{-\Phi/kT}d\mathbf{x}\\&=-kT\ln\prod_{i=1}^{\nu}\int\exp\left(-\frac{x_i^2}{2(kT/\kappa_i)}\right)dx_i\\&=-kT\ln\prod_{i=1}^{\nu}\left(\frac{2\pi kT}{\kappa_i}\right)^{1/2}.\end{aligned} \tag{124}$$

It is clear that $\langle x_i^2\rangle=\sigma_i^{\,2}=kT/\kappa_i$. Now the sum of independent squares of a random variable with a standard Gaussian distribution is a χ^2-variable,[25] so

$$\sum_i^{\nu} \left(\frac{\kappa_i}{kT}\right) x_i^2 = \sum_i^{\nu} \frac{x_i^2}{\sigma_i^2} = \chi^2 \tag{125}$$

and therefore

$$\rho(\chi^2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} (\chi^2)^{\nu/2-1} e^{-\chi^2/2}. \tag{126}$$

Clearly, the χ^2-distribution is a Gamma distribution (see Eq. 67) with $\xi=\chi^2$, $a=\nu/2$ and $\theta=1/2$. Now we can express the distribution of the potential energy Φ, since $\chi^2=2\Phi/kT$, as

$$\rho(\Phi) = \frac{(1/kT)^{\nu/2}}{\Gamma(\nu/2)} \Phi^{\nu/2-1} e^{-\Phi/kT}, \tag{127}$$

which is again a Gamma distribution with $\xi=\Phi$, $a=\nu/2$ and $\theta=1/kT=\beta$. Therefore the harmonic solid has exactly the energy distribution of the Gamma state and using Eq. 65 in Eq. 10 we obtain exactly the expression of the free energy, Eq. 124. The state of this ideal harmonic solid corresponds to the simplest possible solution of a quasi-Gaussian system. However, this Gamma state does not fulfill restriction $R5$, and therefore all derivations starting from Eq. 68 are not valid. Such a system can be regarded to be in a Gamma state with a non converging ideal reduced free energy. In fact, in this case $C_V'=\frac{1}{2}\nu k$, so $(\partial C_V'/\partial T)_V=0$, $b_1=kT$ and hence $\alpha\rightarrow-\infty$. This is a consequence of the fact that we are considering an ideal classical harmonic solid, where the coordinates $\mathbf{x}$ are not bound, and therefore both the configurational volume $\Omega=\int' d\mathbf{x}$ and $\langle e^{\beta\Phi}\rangle$ diverge. It is clear then, that the ratio $\Omega/\langle e^{\beta\Phi}\rangle$ converges to $\Pi_i^{\nu}(2\pi kT/\kappa_i)^{1/2}$, see Eq. 124.

The classical harmonic solid is a good approximation of a real monatomic solid above its Debye temperature Θ_D, say, where $c_V\cong 3R$, the classical limit.[5] For many solids this Debye temperature is in the range of 150–400 K, so in the order of room temperature. Below Θ_D, the previous approach cannot be used, because of the quantum character of the vibrations. In that case the Einstein or Debye approximation can be used.[5]

We also tested the effective Gamma approximation (Eqs. 115–118) and the constant alpha approximation (Eqs. 119–122) in the case of small molecules, on liquid water and methanol. We obtained experimental reduced free energies A_m^r using equilibrium liquid-gas vapor pressures (see Appendix). We used the values of c_V^r, heat of vaporization $\Delta_v H_m$ (with $U_m^r\cong\Delta_v H_m+RT$), and A_m^r at the reference temperature to calculate $\alpha(T_0)$ and δ_0^* and to predict the temperature behavior of A_m^r, U_m^r and c_V^r using this $\alpha(T_0)$ or δ_0^*. It should be noted that the reference temperature T_0 (and the corresponding equilibrium pressure) must be chosen in such a way, that the vapor at T_0 can be regarded as an ideal gas.

For water, the reference condition was defined at $T_0=300$ K and the corresponding equilibrium pressure. We calculated the reference c_V^r from c_P^r,[17] using the relation

$$c_V^r = c_P^r - \frac{\alpha_V^2}{\kappa} T_0 V_m^l(T_0) \tag{128}$$

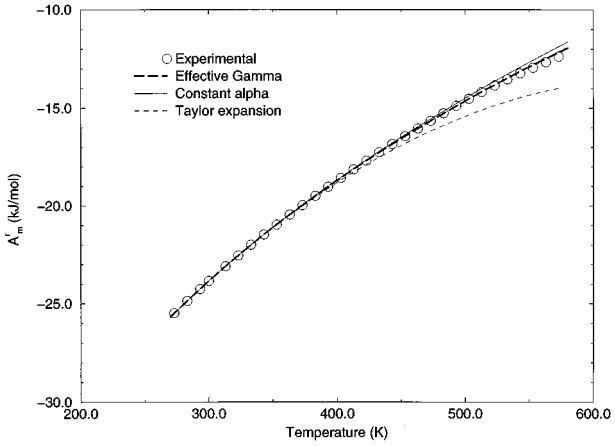

FIG. 1. The free energy A_m^r of liquid water as a function of temperature.

with α_V the isobaric volume expansivity and κ the isothermal compressibility, and $c_P^r=c_P-3R$, correcting for the kinetical part (which is $3R$ for a rigid body). We obtained $c_V^r=0.04942$ kJ/mol K at 300 K. The experimental heat of vaporization, $\Delta_v H_m$ at 300 K was obtained from Atkins,[26] giving $U_m^r=-41.436$ kJ/mol. Values for the equilibrium vapor-pressure were obtained from Schmidt,[27] which are in close agreement with the Handbook of Chemistry and Physics.[17] Above 373 K, say, corrections for the non-ideal behavior of the vapor and the changing of the molar liquid volume V_m^l are not completely negligible. Therefore we corrected all data, as explained in the Appendix, assuming that the deviation from ideal gas is properly described by the second virial coefficient B, and the deviation from an incompressible liquid is described by a linear relation between pressure and molar volume. The values of $B(p^*,T)$ and the isothermal compressibility $\kappa(p^*,T)$ were calculated from the tables in Schmidt.[27] The values of $B(T)$ were in reasonable agreement with those given by other authors.[28] At 573 K, the highest temperature, the correction due to the non-ideal vapor is -1.15 kJ/mol, and compressing the liquid back to the same initial density gives $+0.32$ kJ/mol. Up to 473 K though, both terms almost cancel. Using A_m^r, U_m^r and c_V^r at 300 K, we obtained $\alpha(T_0)=-1.186$ and $\delta_0^*=0.772$.

In Fig. 1 we plot the temperature prediction of the QGE theory using both approximations together with the experimental data. The agreement is very good, even over a temperature range of 300 K. We can see, that there is a slight discrepancy of only 0.5 kJ/mol for the constant alpha and 0.2 kJ/mol for the effective Gamma approximation at the highest temperature. Clearly, the constant alpha approximation starts to become less accurate for such a large temperature interval. In general, we expect α to increase towards the ideal Gaussian state value of $-1/2$ with increasing temperature. An increase of α indeed results in the fact, that the real A_m^r is somewhat lower than the one, predicted with constant α. The prediction of the effective Gamma approximation seems to be more accurate for such a large temperature interval. The agreement of this prediction with experimental data in-

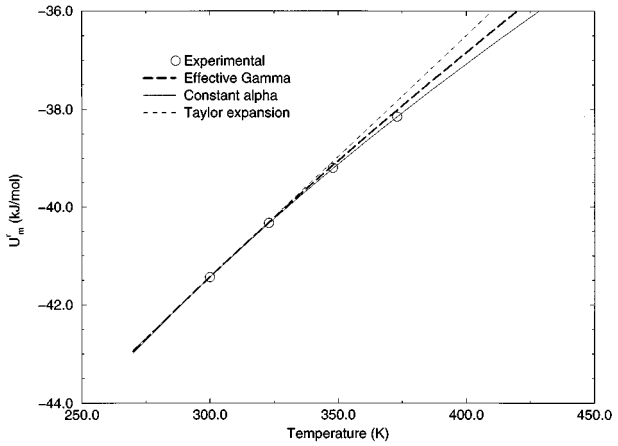

FIG. 2. The potential energy U_m^r of liquid water as a function of temperature.

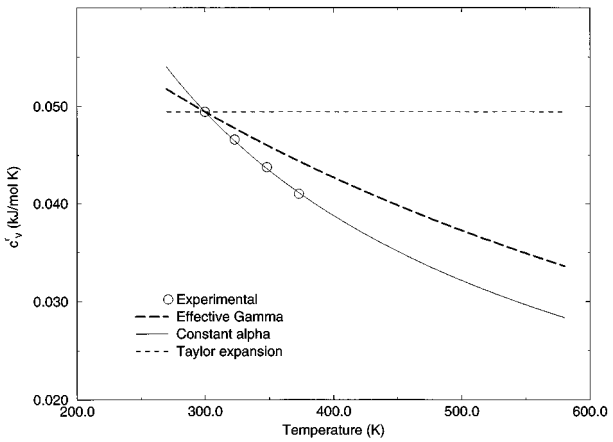

FIG. 3. The heat capacity c_V^r of liquid water as a function of temperature.

dicates that liquid water can be considered as a perturbed Gamma state.

For comparison, we also calculated a second order Taylor expansion of the free energy, using the same amount of experimental data. We can expand A_m^r around T_0 as follows:

$$A_m^r(T)=A_m^r(T_0)+\left(\left.\frac{\partial A_m^r}{\partial T}\right|_{T_0}\right)(T-T_0)+\frac{1}{2}\left(\left.\frac{\partial^2 A_m^r}{\partial T^2}\right|_{T_0}\right)\times(T-T_0)^2 \tag{129}$$

with

$$\left(\left.\frac{\partial A_m^r}{\partial T}\right|_{T_0}\right)=-S_m^r(T_0)=\frac{A_m^r(T_0)-U_m^r(T_0)}{T_0}, \tag{130}$$

$$\left(\left.\frac{\partial^2 A_m^r}{\partial T^2}\right|_{T_0}\right)=\left(\left.\frac{\partial S_m^r}{\partial T}\right|_{T_0}\right)=-\frac{c_V^r(T_0)}{T_0}, \tag{131}$$

where we used $dS_m=dq/T=c_V dT/T$. We see, that the fit of this Taylor expansion is much worse at higher temperatures. This is in a way obvious, since the Taylor expansion does not contain further physical information. It is merely a numerical fit.

This is especially evident, from the comparison of the experimental U_m^r and c_V^r with the prediction of the second order Taylor expansion of the free energy. In the Taylor expansion c_V^r is constant and U_m^r is linear in T. From the specific heat tables at high pressure,[27] we were able to calculate U_m^r at three different temperatures, apart from 300 K, at constant density. Note, that to keep the same density as the equilibrium density at $T_0 = 300$ K, already at 373 K we need to apply a pressure of 908 bar. Results are given in Fig. 2.

From the figure it is evident that $\partial\alpha/\partial T\cong 0$ is an excellent approximation. As already mentioned in Sec. II G the effective Gamma approximation can reproduce with a high accuracy the free energy, but less accurate U_m^r and c_V^r, as it is clear from the figure.

From Schmidt[27] we also obtained values of c_P, $\alpha_V(T)$ and $\kappa(T)$ at the fixed reference liquid density (i.e. at high pressure) in the same temperature range as Fig. 2. Using Eq. 128 we obtained values of $c_V^r(T)$, see Fig. 3.

From these last two figures it is clear, that for U_m^r and c_V^r the constant alpha approximation is excellent and the effective Gamma approximation is still acceptable even for these two properties. This suggests that for liquid water the constant alpha approximation can be considered as an excellent local description, valid for a temperature range of at least 100 K, while the effective Gamma approximation can be considered as a more general description, less accurate for U_m^r and c_V^r, but acceptable over a larger temperature range. Note, that in the range of temperatures between 300 and 400 K the second order Taylor expansion gives a fit to the experimental free energy, which is almost as good as the QGE approximations, while in this range for the other two properties the Taylor expansion is clearly off. This is further evidence, that such an expansion is a pure numerical fit of the free energy only, unable to reproduce other physical properties.

We also applied the theory to liquid methanol. Experimental density and vapor-pressures were obtained from Liley.[29] To calculate c_V^r, we used heat of vaporization values from Liley,[29] but, since we need $\Delta_v H_m$ at constant density, we were forced to calculate c_V^r at low temperature ($T_0 = 180$ K), where the changing in the molar volume is least, obtaining $c_V^r = 0.040$ kJ/mol K and $U_m^r = -40.32$ kJ/mol. For the correction terms at high temperatures, we used values of B from Smith[30] and values of κ at one atmosphere in the range from 273–323 K.[17] For higher temperatures, we used extrapolated data. In this case $\alpha(T_0)=-1.96$ and $\delta_0^*=0.926$.

The results are given in Fig. 4. Also here we have an excellent agreement for both approximations over a large temperature range (300 K). Just as in the case of water, the Taylor expansion becomes quite worse at higher temperature.

Besides the effect of less accurate experimental data and the assumptions, that either $\alpha(T)$ is almost constant or the

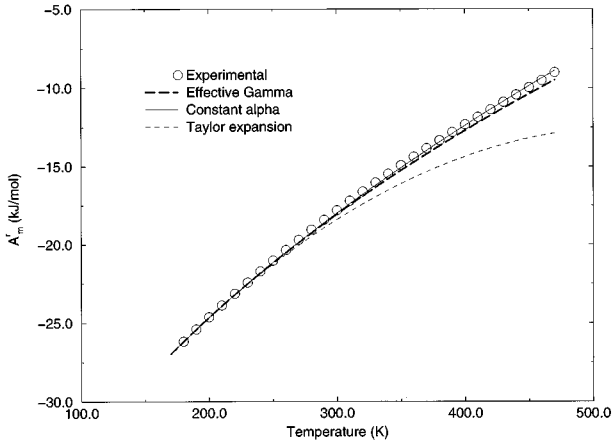

FIG. 4. The free energy A_m^r of liquid methanol as a function of temperature.

statistical state can be considered a perturbed Gamma state, in methanol we have the additional complication of the presence of the dihedral angle H–C–O–H, which we assume to be freely rotating. Anyway, this approximation is within experimental errors, since the experimental c_V of the gas at 180 K is 3.6 R,[29] instead of 3.5 R (rigid body with one dihedral).

IV. DISCUSSION AND CONCLUSIONS

In this article we have shown that rewriting the equilibrium statistical mechanics in terms of the potential energy distribution function it is possible to define different statistical ''states,'' according to the type of distribution. We also showed that for a quasi-Gaussian system such states follow from a generalized Pearson differential equation for uninormal distribution curves. The solutions can be ordered by the number of central potential energy moments, that define the curve, where such moments can be written in terms of temperature derivatives of the average potential energy. Since in general moment M_n is a function only of the first $n-1$ derivatives of the average potential energy, it is clear that such solutions can be ordered according to the number of derivatives as well. This implies that for each solution, the equilibrium physics is given only by the average potential energy and a limited number of derivatives, depending on the type of distribution.

We investigated the properties of the first two possible solutions, which are the Gaussian and Gamma distribution, defining the first two possible statistical states, namely the ideal Gaussian and Gamma state. The first one can be regarded as an ideal or limit condition, but the second one is completely physically acceptable. It is a remarkable fact that the equilibrium physics of systems in this Gamma state can be described completely by the average potential energy and its first two temperature derivatives only. We derived for both states expressions for the ideal reduced entropy and free energy.

We also derived a general expression for the ideal reduced entropy of every quasi-Gaussian system in terms of the heat capacity and the quantity α, the intrinsic entropy function. From this expression, we were also able to obtain a general differential equation (thermodynamic master equation), the solution of which describes the temperature dependence of the free energy, average energy and heat capacity of the system. When the exact statistical state of the system is known, it is possible to solve this equation, as we showed for the Gaussian and Gamma state. Even when such a condition is not present it is still possible to obtain approximated solutions. In this paper we described two possible approximations, the effective Gamma and constant alpha approximations. They both gave very accurate results for the free energy over a large range of temperature, when applied to liquid water and methanol (for these liquids the exact Gamma state is not accurate enough), showing that in such systems the intrinsic entropy function is very temperature insensitive and the statistical states of these liquids can be regarded as perturbed Gamma states. The constant alpha approximation was also able to reproduce very accurately the average energy and heat capacity, at least over a range of 100 K, while the effective Gamma approximation proved to be less accurate in the same range of temperature for the latter properties, although the errors never exceeded $\sim 7\%$.

We also proved that the energy distribution of a classical harmonic solid is exactly a Gamma distribution, showing that already the simplest statistical state of a quasi-Gaussian system can exist. Moreover, already at room temperature many real solids are well described by this harmonic model.

The quasi-Gaussian Entropy theory can also be generalized to mixtures of different components. The expression for the ideal reduced free energy of the total system will be given again by

$$A' = U' - TC'_V \alpha(T) \tag{132}$$

and the partial molar Helmholtz free energy and chemical potential for each component i follow from

$$A'_{mi} = \left(\frac{\partial A'}{\partial n_i} \right)_{p,T,n_{j\neq i}}, \tag{133}$$

$$\mu'_i = \left(\frac{\partial A'}{\partial n_i} \right)_{V,T,n_{j\neq i}}, \tag{134}$$

with n_i the number of moles of component i. In the case of mixtures at infinite dilution the function α is completely determined by the solvent. Currently we are developing the theory for mixtures, according to these equations.

It is also interesting to note that the same basic mathematical approach can be used for direct calculations of the (reduced) chemical potential μ^r, since the latter can be expressed[31] as

$$\mu^r = kT \ln \langle e^{\beta\phi} \rangle = \langle \phi \rangle + kT \ln \langle e^{\beta \Delta\phi} \rangle \tag{135}$$

where ϕ is the intermolecular potential energy of *one* particle interacting with the rest of the system, and $\Delta\phi = \phi - \langle \phi \rangle$. Here we have assumed, for simplicity, that the

intramolecular potential is constant. For the distribution function of $\Delta\phi$, $\rho(\Delta\phi)$, we can set up a similar scheme, as we did for the distribution of the total potential energy ν. There is a major point of difference though between $\rho(y)$ and $\rho(\Delta\phi)$; since $\rho(y)$ applies to a macroscopic system, we can use the central limit theorem to show that $\rho(y)$ must be close to a Gaussian distribution. But the interaction energy ϕ is determined mainly by a (very) limited number of local interactions. Therefore $\rho(\Delta\phi)$ is likely to be more asymmetric than $\rho(y)$.

Preliminary results with Molecular Dynamics simulations show indeed that $\rho(\Delta\phi)$ can range from almost symmetric to very asymmetric, depending on the type of interactions. Short range van der Waals interactions in apolar fluids (for example liquid argon, modeled by a Lennard-Jones (12-6) potential[32]) produce a very asymmetric distribution, wheras a polar liquid like water (modeled by a Coulombic and Lennard-Jones potential in the SPC model[33]) results in an almost Gaussian distribution. The slight asymmetry though is very important, since the right tail of the distribution $\rho(\Delta\phi)$ is multiplied by $\exp(\beta\Delta\phi)$ to calculate $\langle\exp(\beta\Delta\phi)\rangle$. A simple Gaussian distribution was used by Levy and coworkers[34] for calculating the free energy of hydration of ions in solution. Such a distribution corresponds to the simplest solution of the generalized Pearson system.

Our future work will concern both development of the theory and connections with experimental and accurate simulation data.

ACKNOWLEDGMENTS

We like to thank Dr. J. Mavri for carefully reading the manuscript and many stimulating discussions. This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and by the Training and Mobility of Researchers (TMR) Program of the European Community.

APPENDIX: EXPERIMENTAL HELMHOLTZ FREE ENERGY

In this appendix we show how to calculate the Helmholtz free energy for a homogeneous liquid system at fixed volume, from the equilibrium vapor pressure at different temperatures. This method is based on two assumptions. First, since the liquid density has to be fixed at ρ_l^0, we need to change the pressure from the equilibrium pressure p^* to p', to force the density from the equilibrium one $\rho_l(p^*,T)$ back to the reference density $\rho_l(p_0^*,T_0)=\rho_l^0$. We assume, in such a process the pressure to be a linear function of the molar liquid volume (first order approximation). Second, we assume that the behavior of the vapor with density ρ_g can be described adequately by a virial expansion, truncated after the second virial coefficient B:

$$\frac{p}{\rho_g kT}\cong 1+B\rho_g . \tag{A1}$$

With these two assumptions, the reduced free energy can be calculated well above the boiling point of the liquid, but will be less accurate at very high temperatures or gas densities.

We can express the chemical potential of a pure liquid at a temperature T and a pressure p' as:

$$\mu_l(p',T)=\mu_l(p^*,T)+\int_{p^*}^{p'} V_m^l(p,T)dp, \tag{A2}$$

where p^* is the liquid-gas equilibrium pressure at temperature T and V_m^l is the molar volume of the liquid. We can identify p' with the pressure to constrain the density. Since the vapor is in equilibrium with the liquid, we may replace the first term on the right hand side by the gas chemical potential:

$$\mu_l(p',T)=\mu_g^{\ominus}(T)+\int_{p^{\ominus}}^{p^*} V_m^g(p,T)dp+\int_{p^*}^{p'} V_m^l(p,T)dp \tag{A3}$$

with V_m^g the molar volume of the gas, $p^{\ominus}$ some reference pressure, chosen in such a way that the vapor at $p^{\ominus}$ behaves like an ideal gas, and $\mu_g^{\ominus}(T)=\mu_g(p^{\ominus},T)$.

For the second term on the right hand side of Eq. A3, we can rewrite Eq. A1 with $V_m^g=1/\rho_g$ to get an explicit expression for $V_m^g(p,T)$. Integrating we find

$$\begin{aligned}\int_{p^{\ominus}}^{p^*} V_m^g(p,T)dp=&RT\ln\left(\frac{p^*}{p^{\ominus}}\right)\\&+RT\ln\left(\frac{\sqrt{1+\frac{4B}{RT}p^*}-1}{\sqrt{1+\frac{4B}{RT}p^{\ominus}}-1}\right)\\&+RT\left\{\sqrt{1+\frac{4B}{RT}p^*}-\sqrt{1+\frac{4B}{RT}p^{\ominus}}\right\}.\end{aligned} \tag{A4}$$

The third term of Eq. A3 can be approximated by

$$\begin{aligned}\int_{p^*}^{p'} V_m^l(p,T)dp&\cong V_m^l(p',T)(p'-p^*)+\frac{1}{2}|\Delta V_m^l|(p'-p^*)\\&=V_m^l(p',T)(p'-p^*)+\frac{1}{2}\frac{|\Delta V_m^l|\Delta V_m^l}{V_m^l(p^*,T)\kappa(p^*)}\end{aligned} \tag{A5}$$

with $\Delta V_m^l=V_m^l(p^*,T)-V_m^l(p',T)$ and $\kappa=-(1/V_m^l)\times(\partial V_m^l/\partial p)_T$ the isothermal compressibility.

The free energy can be calculated from

$$
\begin{aligned}
A_m(p',T) &= \mu_l(p',T) - p' V_m^l(p',T) \\
&= \mu_g^{\ominus}(T) + RT \ln\left(\frac{p^*}{p^{\ominus}}\right) \\
&\quad + RT \ln\left(\frac{\sqrt{1+\frac{4B}{RT}p^*}-1}{\sqrt{1+\frac{4B}{RT}p^{\ominus}}-1}\right) \\
&\quad + RT\left\{\sqrt{1+\frac{4B}{RT}p^*} - \sqrt{1+\frac{4B}{RT}p^{\ominus}}\right\} \\
&\quad + \frac{1}{2}\frac{|\Delta V_m^l|\Delta V_m^l}{V_m^l(p^*,T)\kappa(p^*)} - p^* V_m^l(p',T). \qquad \text{(A6)}
\end{aligned}
$$

For liquids we can safely neglect the term $p^* V_m^l(p',T)$.

For the ideal gas at the same temperature and density as the liquid, we find:

$$
A_m^0 = \mu_g^{\ominus}(T) + RT \ln\left(\frac{\rho_l^0 RT}{p^{\ominus}}\right) - RT \qquad \text{(A7)}
$$

with $\mu_g^{\ominus}(T)$ the same as in Eq. A3. From the previous equations it follows that the reduced free energy (Eq. 9) is

$$
\begin{aligned}
A_m^r(p',T) &= RT \ln\left(\frac{p^*}{\rho_l^0 RT}\right) + RT \\
&\quad + RT\left\{\ln\left(\frac{\sqrt{1+\frac{4B}{RT}p^*}-1}{\sqrt{1+\frac{4B}{RT}p^{\ominus}}-1}\right)\right. \\
&\quad \left. + \sqrt{1+\frac{4B}{RT}p^*} - \sqrt{1+\frac{4B}{RT}p^{\ominus}}\right\} \\
&\quad + \frac{1}{2}\frac{|\Delta V_m^l|\Delta V_m^l}{V_m^l(p^*,T)\kappa(p^*)}. \qquad \text{(A8)}
\end{aligned}
$$

The last equation provides a very accurate method to evaluate the Helmholtz free energy of a liquid within the range of temperatures where the liquid-gas equilibrium is present. Note, that the first two terms on the right hand side represent the free energy, assuming the vapor to be ideal and the liquid to be incompressible. The third one is the correction term due to the non-ideality of the vapor, and the last term the correction due to the compression of the liquid. Also note, that the last two terms have opposite sign.

[1] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, 3rd ed. (Pergamon, Oxford, 1980), Vol. 5.
[2] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[3] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic, London, 1986).
[4] J. A. Barker and D. Henderson, J. Chem. Phys. **47**, 2856 (1967).
[5] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
[6] J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).
[7] C. G. Gray and K. E. Gubbins, *Theory of Molecular Fluids* (Oxford University, Oxford, 1984).
[8] J. K. Patel, C. H. Kapadia, and D. B. Owen, *Statistics: Textbooks and Monographs* (Marcel Dekker, New York, 1976), Vol. 20.
[9] This condition is not necessarily fulfilled in the case of Molecular Dynamics simulations. In this situation an investigation on the dependence of the energy moments on the size of the periodic box is required.
[10] H. Cramér, *Mathematical Methods of Statistics* (Princeton University, Princeton, 1946).
[11] K. Pearson, Philos. Trans. R. Soc. A **185**, 719 (1894).
[12] K. Pearson, Philos. Trans. R. Soc. A **186**, 343 (1895).
[13] W. P. Elderton, *Frequency-Curves and Correlation* (Layton, London, 1906).
[14] J. K. Ord, *Families of Frequency Distributions* (Griffin, New York, 1972).
[15] N. I. Johnson and S. Kotz, *Continuous Univariate Distributions - 1* (Houghton Mifflin, New York, 1970).
[16] K. A. Dunning and J. N. Hanson, J. Stat. Comp. Simul. **6**, 115 (1978).
[17] *CRC Handbook of Chemistry and Physics*, 75th ed., edited by D. R. Lide (Chemical Rubber, Cleveland, 1994–1995).
[18] N. I. Johnson and S. Kotz, *Discrete Distributions* (Houghton Mifflin, New York, 1969).
[19] K. Tödheide, in *Water: A Comprehensive Treatise*, edited by F. Franks (Plenum, New York, 1972), Vol. 1, Chap. 13, pp. 463–514.
[20] J. P. Phillips, *Spectra-Structure Correlation* (Academic, New York, 1964).
[21] F. R. Dollish, W. G. Fateley, and F. F. Bentley, *Characteristic Raman Frequencies of Organic Compounds* (Wiley, New York, 1974).
[22] J. P. M. Postma, Ph.D. thesis, University of Groningen, The Netherlands, 1985.
[23] C. J. Pouchert, *The Aldrich Library of Infrared Spectra*, 3rd ed. (The Aldrich Chemical Company, Milwaukee, 1981).
[24] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1967).
[25] H. O. Lancaster, *The Chi-Squared Distribution* (Wiley, New York, 1969).
[26] P. W. Atkins, *Physical Chemistry*, 3rd ed. (Oxford University, Oxford, 1987).
[27] E. Schmidt, *Properties of Water and Steam in SI-Units* (Springer, Berlin, 1969).
[28] G. S. Kell, in *Water: A Comprehensive Treatise*, edited by F. Franks (Plenum, New York, 1972), Vol. 1, Chap. 10, pp. 363–412.
[29] P. E. Liley, Chem. Eng. **89**, 59 (1982).
[30] D. S. Smith and R. Srivastava, *Thermodynamic Data for Pure Compounds* (Elsevier, Amsterdam, 1986), Vol. B.
[31] K. S. Shing and K. E. Gubbins, Mol. Phys. **46**, 1109 (1982).
[32] M. P. Allen and D. J. Tildesly, *Computer Simulation of Liquids* (Oxford University, Oxford, 1989).
[33] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
[34] R. M. Levy, M. Belhadj, and D. B. Kitchen, J. Chem. Phys. **95**, 3627 (1991).